Nonlinear optimisation via explicit NRTL model solubility prediction for antisolvent mixture selection in artemisinin crystallisation

Nonlinear optimisation via explicit NRTL model solubility prediction for antisolvent mixture selection in artemisinin crystallisation

Hikaru G. Jolliffe, Samir Diab and Dimitrios I. Gerogiorgis*

Institute for Materials and Processes (IMP), School of Engineering, University of Edinburgh, The King's Buildings, Edinburgh, EH9 3FB, United Kingdom

**Corresponding author: D.Gerogiorgis@ed.ac.uk (+44 131 6517072)*

ABSTRACT

Continuous Pharmaceutical Manufacturing (CPM) has the potential to revolutionise the pharmaceutical industry, with many expected benefits in terms of cost, efficiency and quality. Process modelling and optimisation are valuable methodologies for comparative technoeconomic evaluations: this paper pursues total upstream CPM cost minimisation via two nonlinear optimisation methods. An explicit NRTL solubility estimation method has been used in order to analyse the potential performance of three binary antisolvent mixtures for the crystallisation of artemisinin. This Active Pharmaceutical Ingredient (API) is a key antimalarial substance and has been the focus of several CPM studies, including continuous chemistry and separation.

Ethanol, acetone and ethyl acetate are the three antisolvents studied. Used as a binary antisolvent with toluene as the solvent, the composition of the binary antisolvent, the quantity of the binary antisolvent with respect to the process solvent, and the temperature the overall mixture is cooled to in the crystallisation process are formulated as the three key variables in a nonlinear optimisation problem. Two solvers are used, NOMAD and NLOPT-BOBYQA. Results show that they have very similar precision (<1 % difference), but that the latter is up to 66 % faster; NLOPT-BOBYQA was used for the majority of optimisation cases.

Results show that for the size of the temperature gradient of the crystallisation has a much stronger effect on the total cost than the quantity of antisolvent used. Nearly pure antisolvent use (ethyl acetate followed by ethanol) is favoured as yielding the lowest total cost (for a single crystalliser). Considering multiple units, results indicate that higher API recovery, E-factor, and OpEx benefits can be achieved when using two crystallisers, but metrics are inferior with three or more. For sequential crystallisers of varying size, increasing the number of crystallisers also benefits OpEx, E-factor, and API recovery, but CapEx continues to increase, thereby promising technical but no economic benefits.

KEYWORDS

Continuous Pharmaceutical Manufacturing (CPM); Process modelling; Process optimisation; Nonlinear Programming (NLP); Crystallisation; Artemisinin

1 INTRODUCTION

Given impetus by financial and market pressure, Continuous Pharmaceutical Manufacturing (CPM) is a recent, vibrant field of research aiming to bring the benefits of continuous processing to pharmaceutical production.[1] Traditional batch methods, while allowing equipment flexibility and having extensive expertise in a mature technology, is hampered by disadvantages including low efficiency (material and energy), problematic process scale-up and poor heat and mass transfer. The methodologies of CPM offer methods to adhere to the concept of Quality by Design (QbD) – fully adopted by the pharmaceutical industry in the USA – and is closely related to the development of Process Analytical Technology (PAT), which are novel mechanisms and designs for process monitoring and control. The three fields of CPM, QbD and PAT have approval at the highest regulatory levels.[2]

Artemisinin is one of the most important antimalarial substances available today. First identified and isolated from the plant *Artemisia annua* in the late 1970s,[3] work for which the 2015 Nobel Prize in Physiology or Medicine was awarded, artemisinin is currently produced via batch extraction from the cultivated plant. However, long product lead times and fluctuating demand due to unpredictable burdens of malaria, lead to highly variable prices and production levels.[4] Recent research has demonstrated the continuous synthesis of artemisinin,[5] using a waste product from the current batch production process as a feedstock, producing artemisinin via two sequential plug flow reactors.

In recent years, the continuous flow synthesis of a wide variety of APIs has been demonstrated.[6] However, while there have been significant advances in microscale flow reactor technology, integrated continuous separation technologies are still only gradually becoming available.[7] In addition to experimental work, simulation and cost estimation via process models have also been widely used for evaluating CPM processes, and the field of mathematical optimisation has likewise been successfully used in many CPM studies.[8] A typical example is the nonlinear optimisation of the continuous synthesis of Lorcaserin, which covers many key aspects of CPM and optimisation techniques. The nonlinear formulation includes over 50 reactions and optimizable parameters, and the problem was solved to global optimality with a far fewer number of initial approximations because of the efficient use of the Levenberg-Mardquadt algorithm.[9] Another use of nonlinear optimisation for CPM is the evaluation of the potential benefit from adding a membrane separation operation to a continuous crystallisation process.[10] By using multi-objective optimisation to systematically maximize or minimize critical process parameters within specified ranges, the design space was mapped. The boundaries of attainable product attributes were also found by their maximization/minimization, subject to the range of critical process parameters. Results indicated that the addition of a membrane separation can significantly enlarge the attainable crystal size as well as the design space.[10] In another case, optimisation of stage conditions was used to improve both yields and purity of a continuous MSMPR crystallisation cascade.[11] The solution of population and mass balance models and crystallisation kinetics allowed the estimation of product purity and yield, with results showing that while near-maximum yields can be achieved, it requires multiple crystallizers with quickly diminishing returns.

Continuous downstream product separation and recovery as also attracted research interest in terms of optimisation and control.[12,13] Kriging and Radial Basis-functions (RBF) have been used to for surrogate modelling toward solving black-box feasibility problems.[14–16] Regions of high prediction uncertainty can be efficiently sampled with the proposed RBF method, via adaptive sampling, and the RBF method consistently out-performed a previous method.[15]

Other groups have also focused on the detailed modelling and optimisation of downstream processes. By combining a moving horizon-based optimisation with a hybrid model predictive control, maximum profit for a continuous tableting process that adheres to QbD principles was ensured.[17] Researching the use of nanofiltration membrane cascades for solvent recovery in continuous pharmaceutical processes in a typical nonlinear optimisation problem implemented with the GAMS software package, a group found that total costs could be significantly reduced by implementing solvent recovery.[18,19]

Scheduling of continuous production campaigns has also been explored. After studying the suitability of existing optimisation methods for continuous production lines implemented in a campaign manner, a novel approach was proposed where the objective is on-spec product

maximization as opposed to the commonly used (in higher-volume sectors) start-up/shut-down time minimization.[20] The new approach can be reliably solved with gradient-based methods due to the differentiability of the formulations used.

Life-cycle assessments of CPM processes have also incorporated optimisation as a core component.[21] While systematic optimisation and the consideration of a wide range of performance criteria can reduce Life Cycle Impact Assessment (LCIA) scores by up to 45%,[22] with the emergence of novel technologies, the standardization of methods to address data gaps and sensitivity analyses are required.[23]

Continuous product separation and crystallisation are key operations in the drive towards realizing the implementation of CPM. In previous work, we explored the potential performance of eight antisolvents for continuous crystallisation of artemisinin using both predicted and experimentally reported solubilities.[4,24] We have also previously employed nonlinear optimisation toward finding ideal continuous separation scenarios for ibuprofen: the preferred temperatures, solvent and liquid-liquid extraction configuration for the lowest total design cost for a CPM process were determined while comprehensively considering mass transfer, thermodynamics and economic modelling.[25]

In the present work, we further investigate the continuous separation of artemisinin by formulating a nonlinear optimisation problem considering solubility prediction via an activity coefficient model (NRTL), and environmental impact and sustainability via the *E*-factor, for potential crystallisation binary antisolvent mixtures of ethanol, ethyl acetate and acetone. The CPM process studied here, based on the developed chemistry,[5] has been shown to be compatible with a range of separation methods.[26]

The paper is structured as follows: first, the continuous flow chemistry is presented. Second, the continuous separation of the API product is discussed, also summarizing previous work and how their results are used and inform the present work. This is followed by the elucidation and construction of the nonlinear optimisation problem; key equations regarding process modelling, product separation and economics. The results are then presented, with a discussion of the trends shown and their implications.

2 API PRODUCTION AND SEPARATION

The flowsheet and chemistry studied here is based on the work of Kopetzki and coworkers,[5] illustrated in Figure 1, which we have studied in previous publications.[4,24] The key ingredient is dihydro artemisinic acid (DHAA), which enters a photooxidation reactor (AR-R201) where it is transformed into an intermediate, which in turn undergoes reactions in a second reactor (AR-R202) to ultimately produce the API, artemisinin;[5] the process flowsheet is shown in Figure 1.

Both crystallisation and chromatography have been used experimentally to separate artemisinin from solvent streams:[26,27] simulated moving bed (SMB) chromatography can remove remaining key ingredient DHAA and DCA (which is poorly soluble and so is prone to precipitate in crystallisation operations) while crystallisation can separate a majority of the other impurities. Conceptually illustrated in Figure 1, the process considered here uses a SMB chromatography step with a reported performance,[26] with subsequent solvent removal to achieve the required crystallisation feed concentration of 90% API solubility. For this study, it is assumed that negligible impurities reach the crystallisation step, and that none co-crystallize with the API. The focus here is on the attained quantity of API recoverable in the crystallisation step, estimated via explicit NRTL solubility predictions.

3 PROCESS MODELLING AND NONLINEAR OPTIMISATION

3.1 Reactor and Separation Process Design

Key equations include those for reactor sizing, API solubility, and continuous oscillatory baffled crystallizer (COBC) design. Details for the equations and key variables of reactor volumes (V_{PFRm}, equation 1), reaction rate (r_{mA}, equation 2), energy requirement for crystallisation cooling (q_{COBC}, equation 3) and theoretically attainable API product recovery ($R^{\mathrm{T}}{}_{\mathrm{Crys}}$) are as per our previous publications[25,28] and specifics can be found therein.

$$V_{PFRm} = Q_m C_{mi0} \int_0^{X_{mif}} \frac{dX_i}{-r_{mi}} \tag{1}$$

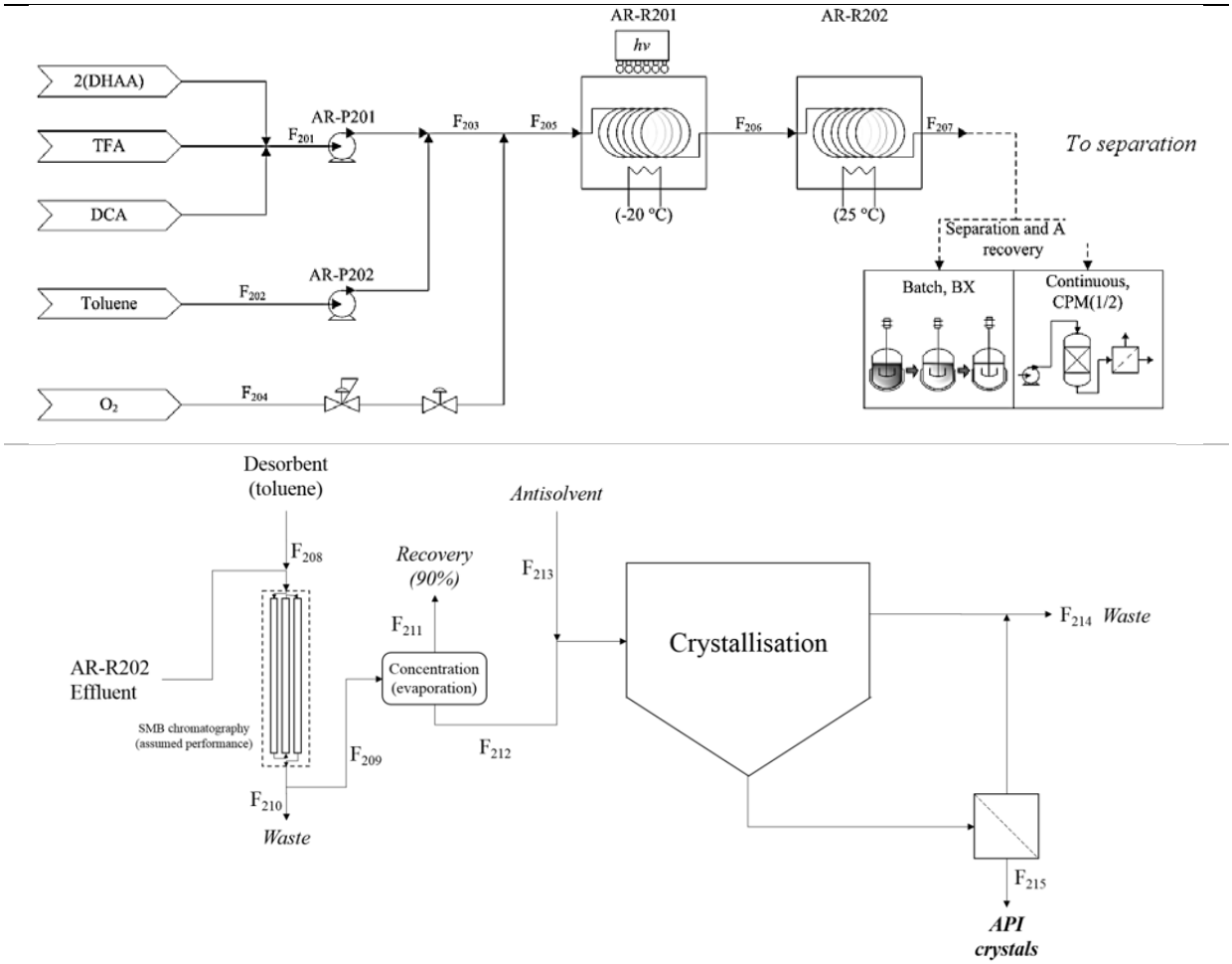

Figure 1. Flowsheet highlighting key steps and processes. Top: API production.[5] Bottom: API separation.[26]

$$r_{mi} = k_{mi} C_{mi} \tag{2}$$

$$q_{COBC} = \left(C_{P,Tol} n_{Tol,COBC} + C_{P,AS} n_{AS,COBC}\right) \Delta T_{COBC} \tag{3}$$

$$A_{COBC} = \frac{q_{COBC}}{U_{COBC} LMTD_{COBC}} \tag{4}$$

$$\dot{m}_{CW} = \frac{q_{COBC}}{C_{P,CW} LMTD_{COBC}} \tag{5}$$

$$R^T_{Crys} = \frac{\left(1 - m_{Crys,API,0}\right)\left(\frac{\left(\dot{m}_S + \dot{m}_{AS}\right) m_{Crys,API,0}}{\left(1 - m_{Crys,API,0}\right)}\right) + \left(\frac{\left(\dot{m}_S + \dot{m}_{AS}\right) m_{Crys,API,f}}{\left(1 - m_{Crys,API,f}\right)}\right)}{\left(\dot{m}_S + \dot{m}_{AS}\right) m_{Crys,API,0}} \tag{6}$$

$$R_{Crys} = f_{Crys} R^T_{Crys} \tag{7}$$

3.2 NRTL Solubility Modelling

In previous work, eight antisolvents were evaluated for crystallisation, with the cases optimized for minimal total cost. Surrogate equations for API solubility were used, based on predictions using the UNIFAC model;[29] ethanol, acetone and ethyl acetate all showed promise. A temperature range of 35 to 5 °C was studied for the crystallisation (the temperature to be cooled to from a feed temperature of 40 °C), for an antisolvent feed rate from 50% to 80% w/w of mother liquor in the crystallizer.

A variety of thermodynamic models exist for the estimation of thermodynamic phase equilibria, including solid-liquid equilibria for the prediction of drug solubilities.[30] The Non-Random Two-Liquid (NRTL)[31,32] and the Non-Random Two-Liquid Segment Activity Coefficient (NRTL-SAC) models[33] are powerful tools for drug solubility modelling,[31] subject to API interaction parameter availability. The present work explicitly estimates solubilities using the original NRTL model,[31,32] for which all required interaction parameters can be determined on the basis of published literature.[34,35]

Here, the solubility of species i, x_i, is computed from the ideal solubility, x_i^{id}, and the activity coefficient γ_i (equation 8). The ideal solubility (at temperature T) is estimated from the enthalpy of fusion ΔH_{fus}; the average difference in heat capacity (ΔC_p) between the temperature and the API melting point, T_{fus}, assumed to be 77.2 J mol^{-1} K^{-1} in equation 9;[34] and R, the universal gas constant. In the NRTL method, the activity coefficient of species i in a n-component mixture, γ_i, is given by equation 10. The parameter α_{ij} is a measure of the non-randomness of the interaction between two species ($\alpha_{ij} = \alpha_{ji}$ and $\alpha_{ii} = \alpha_{jj} = 0$). The dimensionless interaction parameter, τ_{ij}, describing the temperature dependence of solute solubility, can take several forms; one of these forms uses the interaction parameter, b_{ij}, described in equation 12 ($b_{ii} = b_{jj} = 0$).

$$\ln x_i = \ln x_i^{id} - \ln \gamma_i \tag{8}$$

$$\ln x_{id} = \frac{\Delta H_{fus}}{RT}\left(\frac{T}{T_{fus}} - 1\right) - \frac{\Delta C_p}{R}\left(1 - \frac{T_{fus}}{T} + \ln\frac{T_{fus}}{T}\right) \tag{9}$$

$$\ln(\gamma_i) = \frac{\sum_{j=1}^{n} x_j \tau_{ji} G_{ji}}{\sum_{k=1}^{n} x_k G_{ki}} + \sum_{j=1}^{n} \frac{x_j G_{ij}}{\sum_{k=1}^{n} x_k G_{kj}}\left(\tau_{ij} - \frac{\sum_{m=1}^{n} x_m \tau_{mj} G_{mj}}{\sum_{k=1}^{n} x_k G_{ki}}\right) \tag{10}$$

$$G_{ij} = \exp(-\alpha_{ij}\tau_{ij}) \tag{11}$$

$$\tau_{ij} = \frac{b_{ij}}{RT} \tag{12}$$

Published experimental solubility data for artemisinin in pure antisolvents[34] have been used for NRTL model parameter estimation, b_{ij}, regarding API-antisolvent interactions; antisolvent-antisolvent interactions have been taken from the literature.[35] Parameter regression for individual data sets (i.e. binary antisolvent mixture choices) from the experimental data ensured good agreement between predicted API solubilities in antisolvent mixtures from the NRTL model and the published values (Figure 2). In practice, a value of 0.3 is frequently used for $\alpha_{solvent\text{-}solvent}$, and 0.4 is frequently used for $\alpha_{solute\text{-}solvent}$. The required NRTL model parameters which have been determined for study of different systems (binary antisolvent mixtures) via the use of Eqs. (11) and (12) are provided in Table 1.

Table 1. NRTL parameters computed for API solubility prediction in this work.

System	Pairwise Interaction	$\alpha_{12} = \alpha_{21}$	b_{12} (J mol^{-1})	b_{21} (J mol^{-1})	Data Source
API-EtOH-EtOAc	API(1)-EtOH(2)	0.4	3,320.39	3,127.56	(34)
	API(1)-EtOAc(2)	0.3	9,602.49	210.35	(34)
	EtOH(1)-EtOAc(2)	0.3	1,112.11	1,543.52	(35)
API-EtOH-Acetone	API(1)-EtOH(2)	0.4	-229.96	8,062.90	(34)
	API(1)-Acetone(2)	0.3	84.76	6,430.00	(34)

	EtOH(1)-Acetone(2)	0.3	3,905.35	-1,763.81	(35)
	API(1)-Acetone(2)	0.4	-2,391.11	10,248.99	(34)
API-Acetone-EtOAc	API(1)-EtOAc(2)	0.3	-2,068.88	6,014.031	(34)
	Acetone(1)-EtOAc(2)	0.3	-54.14	5,265.98	(35)

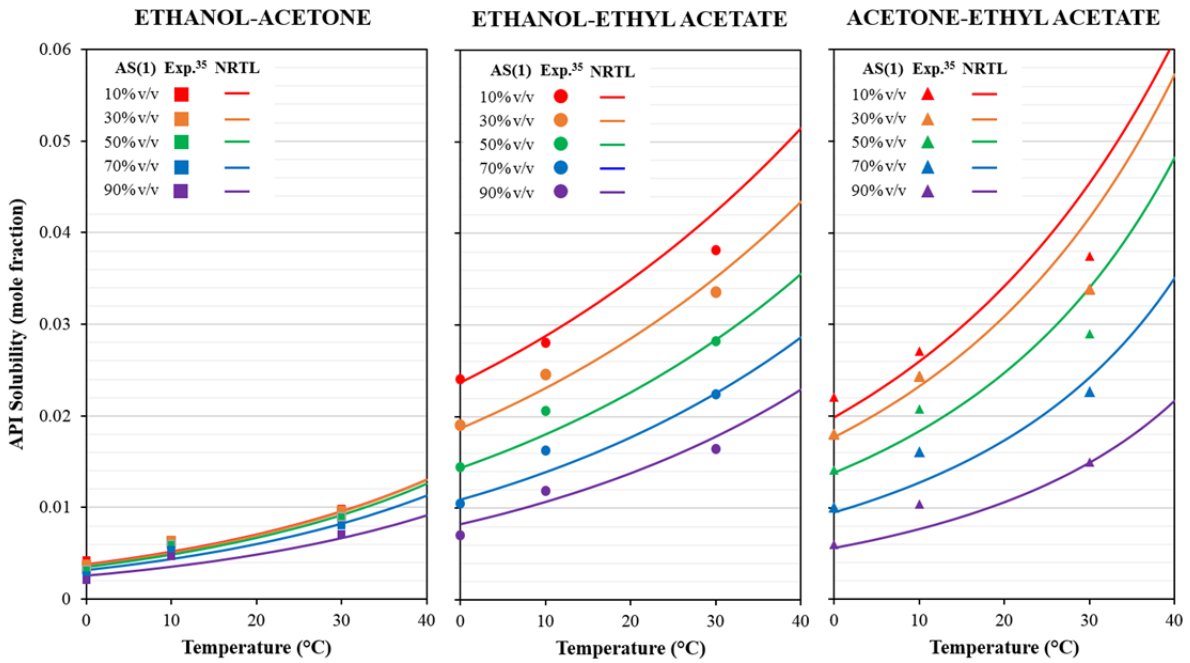

Figure 2. Comparison of reported and NRTL-predicted artemisinin solubility in binary mixtures of ethanol, acetone and ethyl acetate.

In this work, binary antisolvent mixtures of pure antisolvents (ethanol, ethyl acetate and acetone) are considered for API crystallisation; these pure antisolvents are relatively benign.[36] NRTL predictions of API solubilities in different binary antisolvent mixtures under different temperatures and mixture compositions were compared to experimental results for validation. The NRTL code showed good agreement with experimental values[34] (Figure 2). There is some variation with reported values; the small differences for all antisolvent pairs indicate slight under-predictions. Solubilities estimated via NRTL seem somewhat more influenced by temperature (marginally steeper gradients). A non-monotonic solubility trend (esp. in terms of its variation vs. variable mixture composition) would induce nonconvexity, thus significantly perplexing crystallisation optimisation and operability.

3.3 Cost Estimation and Economic Modelling

The cost estimation uses a combination of vendor data and established cost estimation methods to calculate *CapEx* and *OpEx* contributions.[37] The Battery-Limits-Installed-Cost (*BLIC*) (equation 13) estimated via the Chilton Method (Couper, 2003), uses factors for the installed equipment (f_{ins}), process piping (f_{pip}), instrumentation (f_{ins2}), and engineering and construction (f_{ec}) costs, as well as working capital (f_{wc}). The factors used in Eqs. 13–19 are summarized in Table 2. As with previous work,[25,28] factors are also used to estimate the working capital f_{wc} (equation 15) and contingency (equation 16) costs.[37] Where appropriate, vendor prices were used for process equipment, and capacity-cost correlations were used for the remainder.[38]

Reference and design values for equipment are denoted by subscripts *a* and *b*, respectively, while equipment capacities (units for which depend on the equipment type) are denoted by *S* in equation 14, and reference costs by γ_a. Chemical engineering plant cost indices (CEPCIs) were used where

necessary to compute appropriate inflation adjustments. The number of units of equipment i are indicated by n_i. The other parameters in equation 14 are the exponent p_i (an empirical value commonly ranging from 0 to 1 depending on the equipment type) and f (a summary of factors used to account for the effect on cost of various design choices for the equipment such as construction material).

$$BLIC = FOB_{tot} f_{ins} \left(1 + f_{pip} + f_{ins2}\right)\left(1 + f_{ec}\right) \tag{13}$$

$$FOB_{Tot} = \sum_i f_{\gamma,i} n_i \gamma_{ai} \left(\frac{S_{bi}}{S_{ai}}\right)^{p_i} \tag{14}$$

$$\gamma_{wc} = f_{wc} \gamma_{mat} \tag{15}$$

$$\gamma_{con} = f_{con} BLIC \tag{16}$$

$$\gamma_{mat} = \sum_j r_j \pi_j \tag{17}$$

$$\gamma_{utl} = f_{utl} \sum_j R_j \tag{18}$$

$$\gamma_{was} = f_{was} \dot{v}_{was} \tag{19}$$

The cost of material purchase γ_{mat} in equation 17 (chiefly composed of raw ingredients, process solvent and antisolvents) are computed from material prices π_j (sourced from vendors and official shipping records from a range of countries) and the required quantities r_j. Utilities (γ_{util}) and waste disposal (γ_{was}) costs are estimated using factors (Table 2) as per the costing methodology used here.[37] The former is estimated based on total material input, and the latter is based on the quantity of solvent and antisolvent waste (the majority of the waste). Maintenance and reliability costs are not included in this NLP formulation, but are straightforward to implement subject to data and cost factor availability.

Table 2. Cost factors used in Eqs. (13-19).

f_{ins}	= 1.430	f_{ec}	= 0.300	f_{util}	= 0.96 £/kg
f_{pip}	= 0.300	f_{wc}	= 0.035	f_{was}	= 0.35 £/L
f_{ins2}	= 0.120	f_{cont}	= 0.200		

3.4 Nonlinear Optimisation Formulations

The objective function is the sum of the Capital Expenditure (*CapEx*) and the time-discounted Operating Expenditure (*OpEx*). The problem is formulated as a total cost minimization.

$$\min Cost_{Tot} = CapEx + \sum_{t=1}^{t_{life}} \left\{\frac{OpEx}{(1+y)^t}\right\} \tag{20}$$

s.t.

$$CapEx = BLIC + \gamma_{cont} + \gamma_{wc} \tag{21}$$

$$OpEx = \gamma_{mat} + \gamma_{energy} + \gamma_{utl} + \gamma_{was} \tag{22}$$

$$100 \tfrac{\text{kg}}{\text{yr}} \le \dot{m}_{API,prod} \tag{23}$$

$$5^{\circ}\text{C} \le T_{Crys,LO} \le 35^{\circ}\text{C} \tag{24}$$

$$0.5 \le m_{Crys,AS} \le 0.8 \tag{25}$$

$$0 \le x_{AS1} \le 1 \quad (26)$$

$$0 \le x_{AS2} \le 1 \quad (27)$$

$$x_{AS1} + x_{AS2} = 1 \quad (28)$$

$$1 \le n_{COBC} \le 5 \quad (29)$$

The total cost, Eq. (20) covers the purchase, building and running of a given design for the specified period t_{life} (in this case 20 years). The year-to-year discount rate (y) is assumed to be 5%. The components of *CapEx* are the Battery-Limits Installed Cost (*BLIC*), contingency cost (γ_{cont}) and working capital (γ_{wc}). The *OpEx* components are the material purchase costs (γ_{mat}), energy costs (γ_{energy}), utilities costs (γ_{util}), and waste disposal costs (γ_{waste}). A 335-day/8,040-hour working year was assumed.

The constraints in this formulation are a target API production (Eq. 23), maximum and minimum crystallisation cooling temperatures (equation 24) and binary antisolvent mixture weight percentages of the total crystallisation mother liquor (Eq. 25). In addition, in this work there are also constraints on the binary composition of the antisolvent (Eqs. 26-27). Each antisolvent in the binary antisolvent mixture is limited to mole fraction of 0 to 1 (Eq. 28). In addition, the effect of using multiple crystallizers in series (n_{COBC}) was studied, ranging from one to five (Eq. 29). To avoid the need for a mixed integer formulation, each case (varying number of crystallizers) was solved separately.

The key decision variables in this formulation are the temperature to which the crystallisation operation is cooled ($T_{cryst,LO}$), the antisolvent mixture weight percentage in the mother liquor ($m_{Crys,AS}$), and the composition of the binary antisolvent mixture (x_{AS1}). To allow a comparison to previous work, $T_{cryst,LO}$ and $m_{Crys,AS}$ were restricted to a range of 5 to 35 °C and 50% to 80% for crystallisation cooling temperature and antisolvent weight percent in the mother liquor, respectively.

3.5 NLP code and structure

The optimisation is coded in three portions; a section implementing reactions and reactor sizing, thermodynamics and product recovery equations (Eqs. 1-12); a section implementing cost estimation equations (Eqs. 13-19); and a section implementing the optimisation itself (Eqs. 20-29). Two solvers were tested: the NOMAD solver and the NLOPT solver with the BOBYQA algorithm (NLOPT-BOBYQA), both implemented in MATLAB.

3.5.1 *NLOPT-BOBYQA solver*

The NLOPT solver is a collection of over 20 nonlinear program algorithms, and includes those for both local and global optimisation. It also includes several solvers for derivative-free optimisation, one of which – BOBYQA – is used here, derived from the BOBYQA subroutine and modified to work with NLOPT.[39] BOBYQA performs derivative-free, bound-constrained optimisation using an iteratively constructed quadratic approximation for the objective function; as such, it may perform poorly for objective functions that are not twice-differentiable.

3.5.2 *NOMAD solver*

The NOMAD solver uses a Mesh Adaptive Direct Search (MADS) algorithm to solve non-differentiable and global nonlinear programs and allows rapid convergence for derivative-free optimisation.[40,41] The MADS algorithm generates a trial point on a mesh and iterates to improve the current best solution by initiating the following trial point on a finer mesh. Each iteration is composed of a search and poll step.[40] The search step allows the creation of trial points anywhere on the mesh. The poll step explores the mesh site near the current iterate with different trial points.

3.5.3 *Solver implementation*

Solvers were initiated from a variety of starting locations to check the effect on the optimal solution (visualization of response surfaces indicated that there would not be multiple optima). The starting locations were a grid of initial guesses for binary antisolvent mixture-to-feed ratio ($m_{Crys,AS,0}$, between 55 and 75 % w/w of mother liquor in increments of 10), crystallisation cooling temperature ($T_{Crys,LO,0}$, between 10 °C and 30 °C, in increments of 10) and balance of pure antisolvents in the binary antisolvent mixture ($x_{AS1,0}$ = 25, 50 and 75 mol%). Solver and subroutine tolerances were set at 10^{-9}.

Solution time is dependent on the solver selected. Studying multiple initial guesses (55, 65, or 75% w/w of binary antisolvent mixture in the mother liquor ($m_{Crys,AS}$); 10, 20 or 30 °C for crystallisation cooling temperature ($T_{Crys,\ LO}$); 25, 50 or 75 mol% of the first antisolvent in the binary antisolvent mixture (x_{AS1})) and using similar solver settings as much as possible, the NLOPT solver, using BOBYQA algorithm, is faster than the NOMAD solver for all antisolvent pairs, as shown in Table 3. Further details on the performance and optimisation paths of both implemented solvers are provided in Appendix A. As NLOPT-BOBYQA solves faster than NOMAD without significant differences in precision, it was chosen as the main solver. Except for the solution time, there was no difference from using different initial guesses, and so for most cases an initial guess of x_0 = (55, 50, 10), was used.

Table 3. Solver performance summary for initial guess x_0 = (55, 50, 10) = ($100m_{crys,AS}$, x_{ASr}, $T_{crys,LO}$) for one crystallizer.

	Solution time (s)	
Antisolvent pair	**NLOPT-BOBYQA**	**NOMAD**
Ethanol-Acetone	154	272
Ethanol-Ethyl Acetate	167	160
Acetone-Ethyl Acetate	145	235

4 RESULTS AND DISCUSSION

4.1 Effect of temperature and antisolvent quantity and composition

Total cost response surfaces were produced to illustrate the design spaces of different process configurations and to ensure no local optima were present. Figure 3 shows total cost response surfaces for one implemented crystallizer for all binary antisolvent mixtures considered. Total cost response surfaces for further process configurations are provided separately in Appendix B.

Surfaces in Figure 3 for ethanol-acetone and ethanol-ethyl acetate are nearly identical, as both cases are pushed to >99.8% ethanol in the binary antisolvent mixture used. When one crystallizer is implemented, the optimal binary antisolvent mixture composition for acetone-ethyl acetate usage is >98% ethyl acetate. These results show that nearly pure ethanol is the best candidate antisolvent, followed by ethyl acetate, followed by acetone. The surface plots also show that acetone-ethyl acetate usage for antisolvent crystallisation allows the lowest total costs for one crystallizer.

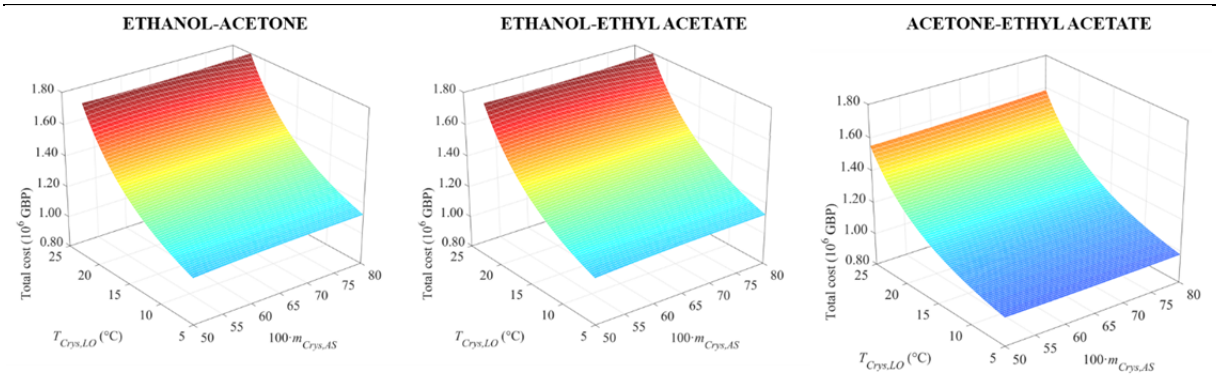

Figure 3: Total cost response surface for one implemented crystallizer.

In all process configurations, the optimal solution (process configuration) is pushed to nearly pure antisolvents (>99.8% ethanol for ethanol-ethyl acetate and ethanol-acetone antisolvent mixtures and >98% ethyl acetate for the acetone-ethyl acetate mixture). Differences in optimal antisolvent mixture compositions influence the minimum total cost attainable, the extent of which depends on the exact composition and the price of the carrier and antisolvent(s) present. Published experimental data shows that artemisinin solubilities increase in the order ethanol < acetone < ethyl acetate. Our NLP optimisation results show that nearly pure antisolvents (ethanol and ethyl acetate, at >99.8% and >98%, respectively), rather than antisolvent mixtures, offer the best economic performance (total cost minimization). Beyond the possible effect of the process carrier solvent (toluene) on solid-liquid

equilibria, it is clear that total cost surfaces (Fig. 3) demonstrate very little sensitivity to the extent of antisolvent addition; while critical for technical operation, the latter has a much weaker effect on economic performance in comparison to cooling temperature, and may explain why pure antisolvents are favoured by the NLP optimisation formulation.

Total costs obtained at maximum and minimum values of two of the key decision variables, binary antisolvent mixture content of the mother liquor ($m_{Crys,AS}$) and cooling temperature ($T_{Crys,LO}$), have been compared to the coordinates corresponding to total cost minima. These values (calculated from the total cost minimization illustrated in the response surfaces in Appendix B) are compared in Figure 4. The histogram shows the total costs for different values of binary antisolvent mixture content of mother liquors for 1-3 fixed and variable crystallizer sizes. Plots corresponding to binary antisolvent mixtures containing ethanol are identical for both fixed and variable crystallizer sizes, as both mixtures are pushed to nearly pure (>99.8%) ethanol with respect to binary antisolvent mixture composition.

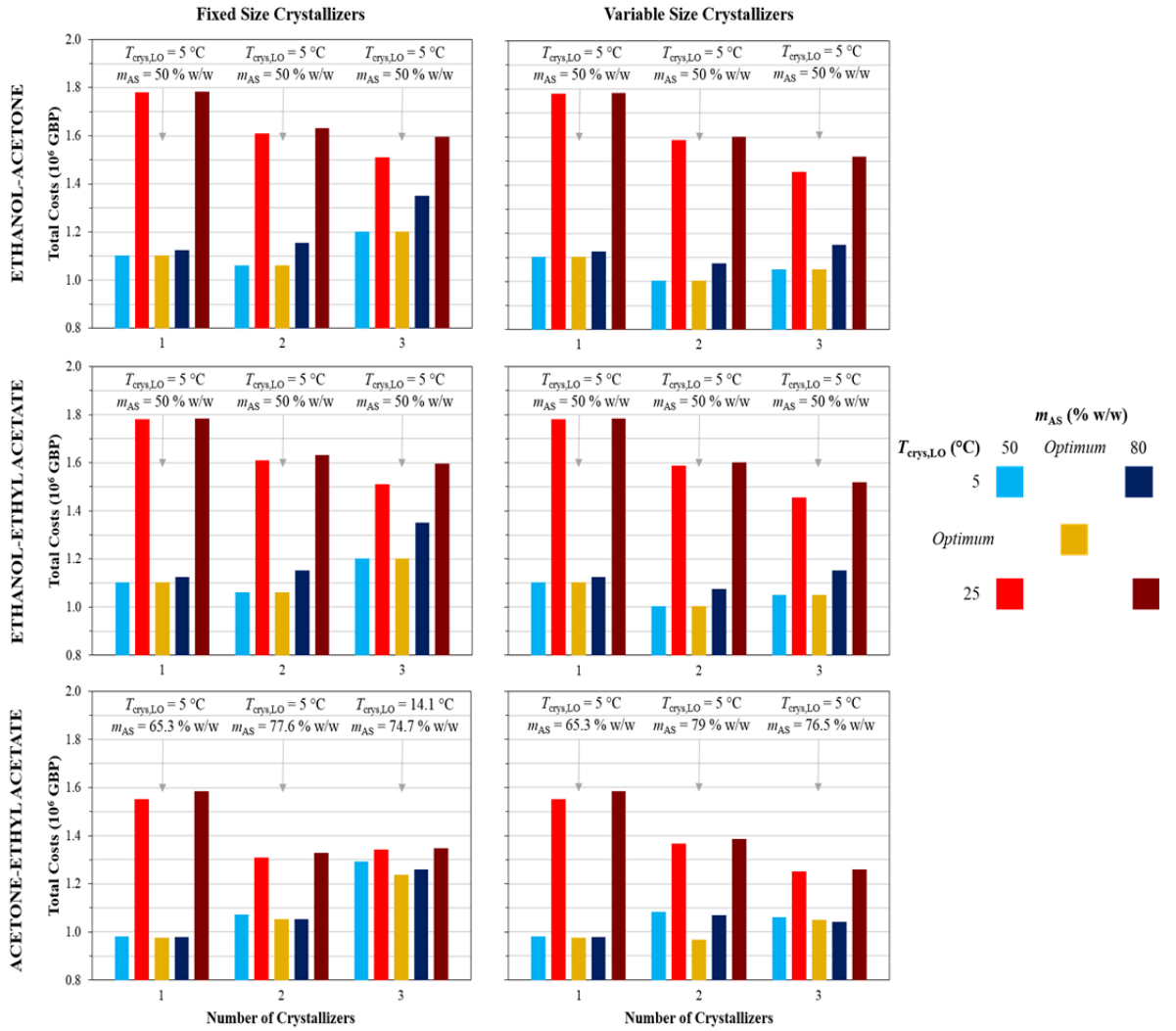

Figure 4: Effect of temperature and binary antisolvent mixture content on total crystallisation system cost, for fixed and variable crystallizer sizes. Optimum crystallisation cooling temperatures (T_{opt}) at antisolvent content, $m_{AS,cryst}$ = 50%.

Generally, a decrease in crystallisation cooling temperature (i.e. lowering the temperature the mother liquor is cooled to) reduces the total cost via an increase in API recovery. Increasing recovery implies less material throughput, hence smaller equipment, lowering costs. For certain instances, the cooling temperature is not pushed to the lower bound of 5 °C as shown by optimal temperatures greater than 5 °C. Only small increases in API recovery are attainable beyond a certain crystallisation cooling temperature, which incurs unnecessary cost increases. If the problem were formulated as a

profit maximization (e.g. including revenue from artemisinin sales) then it is possible that the crystallisation cooling temperature would be pushed to bounds and greater API recovery achieved.

For both binary antisolvent mixtures containing ethanol, the optimal quantity of binary antisolvent mixture was 50% w/w of the mother liquor. This result is different to our previously estimated performances,[4] which is likely due to the greater suitability of the NRTL method over the UNIFAC method.[42] When an acetone-ethyl acetate binary antisolvent mixture is used, the optimal quantity of binary antisolvent mixture varies between 65.3-79.0% w/w of the mother liquor for fixed- and variable-sized crystallizers. While the effect of binary antisolvent mixture quantity (with respect to mother liquor) on total cost is not as pronounced as with crystallisation temperature, i.e. there are steeper gradients in total costs along temperature axes compared to binary antisolvent mixture quantity axes (see Appendix B), the cause of the effect is the same; greater API recoveries allow for a reduction in cost for a given process output.

Total costs improve with acetone-ethyl acetate usage (the optimal composition of which is >98% ethyl acetate for all cases) over ethanol-acetone and ethanol-ethyl acetate (both of which are pushed to bounds of >99.8% ethanol). Nearly pure antisolvent (of those considered) is preferred to binary antisolvent mixture usage. Certainly, with one crystallizer under the performance metrics considered, ethyl acetate usage clearly emerges as the best choice, followed by ethanol, followed by acetone.

4.2 Effect of number and type of crystallizer

Sequential crystallizers of fixed (identical) size and sequential crystallizers of variable (non-identical) size are considered. Commercially, it is very common to use multiple crystallizers in series to improve product recovery, as one single unit operation is often insufficient for the desired performance. The sizes of the crystallizers were determined by the optimisation formulation (Eq. 19). Specifically, it is determined by the required heat transfer area to cool the stream from 40 °C to the cooling temperature (ranging from 5 to 35 °C, Eq. 5). When one crystallizer is considered, solutions for fixed- and variable-size crystallizers are identical.

Figure 5 shows the effect of multiple crystallizer usage on *CapEx*, *OpEx*, recovery and the E-factor on minimum total costs for each process option. The E-factor is a green chemistry metric, and its simplest definition is the mass of waste generated per unit mass of product.[43] *E*-factor values can be as high as 200 for batch-dominated industries such as pharmaceuticals, and as low as 0.1 for continuous production-dominated industries such as oil and gas.[44] In the optimisation cases studied here, the *E*-factor is defined and calculated as the ratio of the mass of waste (consists of byproducts (bpd), unconverted reactants (ur), waste solvent (wS, 10% of total requirements assumed unrecovered), waste antisolvent (wAS, all assumed unrecovered) and unrecovered API (uAPI)) to the mass of pure crystallized API product.

$$E\text{-}factor = \frac{m_{waste}}{m_{API}} = \frac{m_{bpd} + m_{ur} + m_{wS} + m_{wAS} + m_{uAPI}}{m_{API}} \tag{30}$$

As shown in Figure 5, there are notable differences between using crystallizers of fixed- and variable-size, and some similarities. For both types, *CapEx* forms the bulk of the total cost, and rises relatively linearly with the number of crystallizers. For all binary antisolvent mixture cases using fixed-size crystallizers, API recovery notably increases when implementing two crystallizers, but only incremental recovery increases are attainable when three or more crystallizers are used. *CapEx* reaches a minimum at two fixed-size crystallizers for binary antisolvent mixtures containing ethanol and for one fixed-size crystallizer for the acetone-ethyl acetate mixture; *OpEx* reaches a minimum when three fixed-size crystallizers are implemented for all binary antisolvent mixtures. These trends are a consequence of the plateaus reached in API recovery with increasing number of crystallizers.

For crystallizers of variable size, *CapEx* continues to increase, as the costs of the increased number of crystallizers outweighs the cost reductions from smaller reactors (due to less material throughput being needed due to increased API recoveries). API recoveries increase steadily with increasing number of crystallizers; as a result, *OpEx* continually decreases due to less material throughput required, lower material, utilities and waste handling costs are realized. For all binary antisolvent

mixtures cases and both fixed- and variable sized crystallizer implementation, minimum total costs are attained when two crystallizers are implemented. The decreasing *OpEx* is of a smaller magnitude than the increase in *CapEx*, and so the total costs increase with the number of crystallizers.

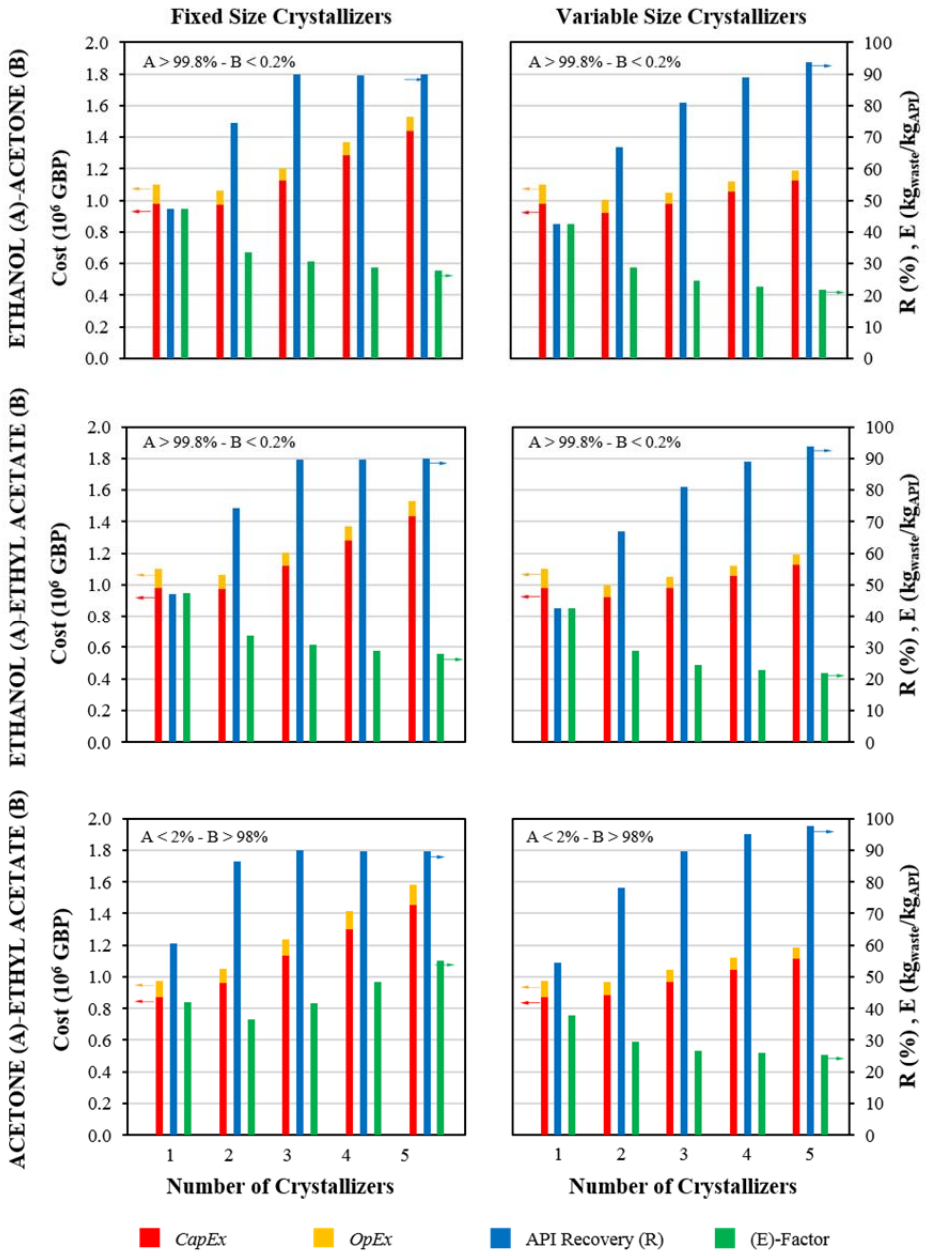

Figure 5. Effect of multiple crystallizer use on total cost, *CapEx*, *OpEx*, API recovery (*R*) and E-factor.

Minimum cost is significantly dictated by how many crystallizers are implemented and whether they are of fixed- or variable-size. The effect of multiple crystallizer implementation for fixed- and variable-size crystallizer volumes is illustrated in Figure 6. For fixed-size crystallizers, total crystallizer volumes increase with the number of implemented crystallizers, whilst the total crystallizer volume for variable-size crystallizers remains relatively constant. Cases using acetone-ethyl acetate antisolvent mixtures (>98% ethyl acetate) require larger total crystallizer volumes than ethanol-ethyl acetate and ethanol-acetone mixtures (both driven to >99% ethanol), however the former still achieves lower total costs than the latter. The effects of improving API recovery (Figure 5) on both *CapEx* and *OpEx* upstream of the separation train, i.e. on allowing smaller reactor volumes

and material requirements caused by the improved separation efficiency, have a larger impact than the required total crystallizer volume.

As this problem has been formulated as a total cost minimization, neither API recovery nor E-factor are prioritized during optimisation. For profit or API recovery maximization, recoveries would continue to improve with additional crystallizer use, at the expense of greater cost (more than the cost increase in Fig. 5).

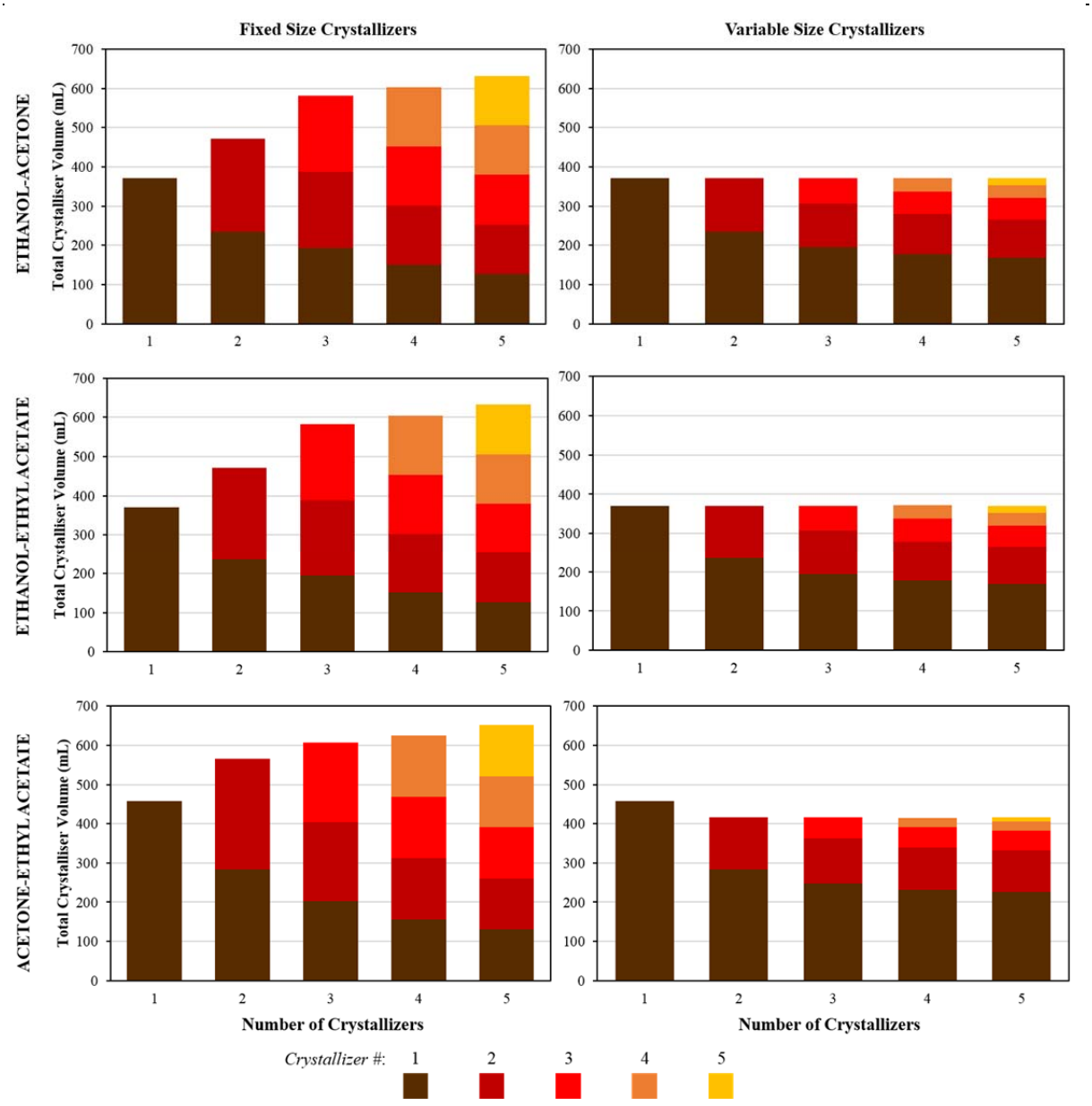

Figure 6: Effect of multiple crystallizer use on individual and total crystallizer volumes.

4.3 Cost and API-recovery/E-factor trade-off

Figure 5 illustrates a trade-off between cost, API recovery and E-factor. For fixed-size crystallizers, there is a modest improvement in E-factor from using two crystallizers instead of one. With more than two crystallizers, E-factors begin to steadily worsen (increase). For two crystallizers, API recovery plateaus, beyond which additional recovery gain is materially inefficient (i.e. worse E-factor values).

For variable-size crystallizers, the E-factor continues to improve (decrease) with the number of crystallizers, as does API recovery, although the greatest improvements are from increasing the number of crystallizers to two or three, with additional units yielding only incremental improvements.

This benefit is not sufficient to outweigh the costs of additional crystallizers. The estimated capital costs are such that multiple crystallizers dominate cost reductions elsewhere. However, the cost penalty from using multiple crystallizers is not as steep with variable-size crystallizers as it is for those of fixed-size (Figure 5). Nevertheless, it is entirely likely that the improvement in API recovery and E-factor justifies a small increase in total cost.

5 CONCLUSIONS

This study conducts nonlinear optimisation for total cost minimization of the steady-state continuous pharmaceutical manufacturing of artemisinin. The problem formulation implements explicit NRTL solubility calculations for artemisinin crystallisation in candidate binary antisolvent mixtures of ethanol-ethyl acetate, ethanol-acetone and acetone-ethyl acetate. Results show that a binary antisolvent mixture of acetone-ethyl acetate (>98% ethyl acetate) attains the lowest total cost of $0.967 \cdot 10^6$ GBP when using two sequential crystallizers of 284 mL and 133 mL, respectively, with an API recovery of 78.1% (crystalline artemisinin) and an *E*-factor of 29.6. Both ethanol-ethyl acetate and ethanol-acetone mixtures are pushed to bounds of >99.8% ethanol.

Studying the use of multiple crystallizers showed that sequential crystallizers allow improved API recovery, *E*-factor, and *OpEx* benefits with the use of two crystallizers, but worsen with three or more. *CapEx* increases with the number of implemented crystallizers in all configurations. For sequential crystallizers of varying size, increasing the number of crystallizers monotonically benefits *OpEx*, *E*-factor, and API recovery. Crystallizers of varying size are preferable over those of fixed size; (nearly) pure ethyl acetate usage attains the lowest total costs. These results illustrate the importance of nonlinear optimisation studies prior to further experimental investigation and scaleup, and informs the development of CPM processes for this societally important antimalarial.

6 AUTHOR INFORMATION

Corresponding Author
*Email: D.Gerogiorgis@ed.ac.uk, phone +44 131 6517072

ORCID
Dimitrios I. Gerogiorgis 0000-0002-2210-6784

Notes
The authors declare no competing financial interest.

7 ACKNOWLEDGEMENTS

The authors gratefully acknowledge the financial support of the Engineering and Physical Sciences Research Council (EPSRC) via Doctoral Training Partnership (DTP) PhD fellowships awarded to Mr. Hikaru G. Jolliffe (Grant No. EP/M506515/1) and Mr. Samir Diab (Grant No. EP/N509644/1). Tabulated and cited literature data suffice for reproduction of all original process simulation results, and no other supporting research data are stored or required.

8 NOMENCLATURE AND ACRONYMS

α	Non-randomness parameter in the NRTL model
API	Active Pharmaceutical Ingredient
b_{ij}	Parameter of the NRTL model
BLIC	Battery-Limits-Installed-Cost of all equipment in a process, part of CapEx, GBP
C_{ai}	Reference cost of item *a*, GBP
CapEx	Capital expenditure, GBP
c_{fi}	Product of factors that consider specifics (e.g. material of construction) of item *i*
C_{mi}	Concentration of species *i* in reactor *m*, mol L^{-1}
C_{mi0}	Initial concentration of species *i* in PFR *m*, mol L^{-1}
CEPCI	Chemical engineering plant cost index

COBC	Continuous oscillatory baffled crystallizer
$Cost_{Tot}$	Total costs over a certain plant lifetime, adjusted to the present, GBP
$C_{p,i}$	Standard heat capacity of component *i* (J mol^{-1} K^{-1})
CPM	Continuous Pharmaceutical Manufacturing
E-factor	Environmental factor, a measure of material efficiency, kg (waste)/kg (product API)
$f_{\gamma,i}$	Factors accounting for different design choices between equipment of type *i*
f_{cont}	Factor relating the contingency to the Battery-Limits-Installed-Cost
f_{crys}	Factor to account for non-attainment of equilibrium in COBC
f_{ec}	Factor relating to engineering and construction costs
f_{ins}	Factor relating to equipment installation
FOB_{Tot}	(Total) Free-on-Board cost of equipment
f_{ins}	Factor relating to installed equipment costs
f_{ins2}	Factor relating to process instrumentation costs
f_{pip}	Factor relating to process piping
f_{utl}	Factor relating the cost of utilities for the total annual material requirements (kg)
f_{was}	Factor relating the cost of waste handling to the volume of waste
f_{wc}	Factor relating the working capital to the cost of materials
G_{ij}	Parameter measuring interaction energy between components *i* and *j* in the NRTL model
ΔH_{fus}	API enthalpy of fusion, J mol^{-1}
k_{mi}	Rate constant of species *i*, reactor *m*, L mol^{-1} hr^{-1}
LCIA	Life cycle impact assessment
LMTD	Log mean temperature difference
$\dot{m}_{API,prod}$	Mass of API produced after separation, kg yr^{-1}
m_{bpd}	Mass of by-product waste, kg yr^{-1}
$m_{Crys,AS}$	Antisolvent weight percent of crystallisation mother liquor, %
$m_{Crys,API,0}$	Initial API mole fraction in COBC
$m_{Crys,API,f}$	Final API mole fraction in COBC
$\dot{m}_i$	Mass flowrate of component *i*, kg s^{-1}
m_{uAPI}	Mass of unrecovered API, kg yr^{-1}
m_{ur}	Mass of unreacted reagent waste, kg yr^{-1}
m_{wAS}	Total mass of waste, kg yr^{-1}
m_{wS}	Mass of solvent waste, kg yr^{-1}
MADS	Mesh Adaptive Direct Search
n_{COBC}	Number of COBCs implemented
n_i	Number of units of equipment *i*
$n_{i,COBC}$	Molar flowrate of component *i* in the crystallizer (mol s^{-1})
OpEx	Operating expenditure, GBP
PAT	Process Analytical Technology
p_i	Empirical value in Eq. 14, varying for different equipment
π_i	Unit price of material *j* (GBP kg^{-1})
QBD	Quality by Design
Q_m	Volumetric flowrate through reactor *m* (mL hr^{-1})
q_{COBC}	Cooling duty of COBC (W)
r_i	Reaction rate of molecule *i*, mol L^{-1} hr^{-1}
R	Universal gas constant (= 8.314 J mol^{-1} K^{-1})
R_j	Annual requirements of material *j*, kg y^{-1}
R^T_{cryst}	Theoretical crystallisation recovery of API at equilibrium, %
R_{cryst}	Actual crystallisation recovery of API (%)
RBF	Radial Basis function
S_{ai}	Reference capacity of item *i*, units depend on item type
S_{bi}	Design capacity of item *i*, units depend on item type
SMB	Simulated Moving Bed
T	Temperature, °C or K
t_{life}	Plant lifetime, years

ΔT_{crys}	Temperature change in COBC, K
$T_{Crys,LO}$	Cooling temperature of crystallisation, °C
T_{fus}	API melting temperature, °C or K
U_{COBC}	Overall heat transfer coefficient in COBC, W m^{-2} K^{-1}
V_{PFRm}	Volume of plug flor reactor *m*, Ml
x_{ASi}	Mole fraction of pure antisolvent *i* in binary antisolvent mixture
x_i	Solubility mole fraction of component *i*
x_i^{ID}	Ideal solubility mole fraction of component *i*
x_{ki}	Mole fraction of species *i* in stream *k*
X_{mi}	Conversion of species *i* in PFR *m*
X_{mif}	Final conversion of species *i* in PFR *m*
x_0	Vector of initial guesses ($100m_{crys,AS}$, x_{ASr}, $T_{crys,LO}$)
y	Discount rate, the correction factor for adjusting costs to the present (%)
γ_{ai}	Reference cost of item *i*, GBP
γ_{con}	Cost of contingency, GBP
γ_{ec}	Factor to account for equipment engineering and construction
γ_{pip}	Factor to account for equipment piping
γ_{con}	Cost of contingency, GBP
γ_i	Activity coefficient of species *i*
γ_{ins}	Factor to account for equipment installation
γ_{ins2}	Factor to account for equipment instrumentation
γ_{mat}	Cost of material raw material purchase (including solvents and catalysts), GBP
$\gamma_{price,j}$	Cost of purchase of material *j*, GBP kg^{-1}
γ_{util}	Cost of utilities, GBP
γ_{was}	Cost of waste handling, GBP
γ_{wc}	Cost of working capital, GBP
$\dot{v}_{was}$	Annual generation of waste (L y^{-1})
τ_{ij}	NRTL model dimensionless interaction parameter

9 REFERENCES

1. Gutmann, B., Cantillo, D. & Kappe, C. O. Continuous-Flow Technology: A Tool for the Safe Manufacturing of Active Pharmaceutical Ingredients. *Angew. Chemie-International Ed.* **54,** 6688–6728 (2015).
2. Lee, S. L. *et al.* Modernizing Pharmaceutical Manufacturing: from Batch to Continuous Production. *J. Pharm. Innov.* **10,** 191–199 (2015).
3. Tu, Y. The discovery of artemisinin (qinghaosu) and gifts from Chinese medicine. *Nat. Med.* **17,** 1217–1220 (2011).
4. Jolliffe, H. G. & Gerogiorgis, D. I. Process Modelling and Simulation for Continuous Pharmaceutical Manufacturing of Artemisinin. *Chem. Eng. Res. Des.* **112,** 310–325 (2016).
5. Kopetzki, D., Levesque, F. & Seeberger, P. H. A Continuous-Flow Process for the Synthesis of Artemisinin. *Chem. Eur. J.* **19,** 5450–5456 (2013).
6. Bana, P. *et al.* The route from problem to solution in multistep continuous flow synthesis of pharmaceutical compounds. *Bioorganic and Medicinal Chemistry* (2016). doi:10.1016/j.bmc.2016.12.046
7. Jensen, K. F. Flow chemistry – Microreaction technology comes of age . *AIChE J.* (2017). doi:10.1002/aic.
8. Escotet-Espinoza, M. S., Rogers, A. & Ierapetritou, M. G. in 281–309 (Humana Press, New York, NY, 2016). doi:10.1007/978-1-4939-2996-2_9
9. Grom, M., Stavber, G., Drnovšek, P. & Likozar, B. Modelling chemical kinetics of a complex reaction network of active pharmaceutical ingredient (API) synthesis with process optimisation for benzazepine heterocyclic compound. *Chem. Eng. J.* **283,** 703–716 (2016).
10. Wang, J. & Lakerveld, R. Continuous Membrane-Assisted Crystallisation To Increase the Attainable Product Quality of Pharmaceuticals and Design Space for Operation. *Ind. Eng. Chem. Res.* **56,** 5705–5714 (2017).

11. Li, J., Lai, T. C., Trout, B. L. & Myerson, A. S. Continuous Crystallisation of Cyclosporine: the Effect of Operating Conditions on Yield and Purity. *Cryst. Growth Des.* **17,** 1000–1007 (2017).
12. Ierapetritou, M., Muzzio, F. & Reklaitis, G. Perspectives on the continuous manufacturing of powder-based pharmaceutical processes. *AIChE J.* **62,** 1846–1862 (2016).
13. Haas, N. T., Ierapetritou, M. & Singh, R. Advanced Model Predictive Feedforward/Feedback Control of a Tablet Press. *J. Pharm. Innov.* **12,** 110–123 (2017).
14. Boukouvala, F. & Ierapetritou, M. G. Surrogate-based optimisation of expensive flowsheet modelling for continuous pharmaceutical manufacturing. *J. Pharm. Innov.* **8,** 131–145 (2013).
15. Wang, Z. & Ierapetritou, M. A novel feasibility analysis method for black-box processes using a radial basis function adaptive sampling approach. *AIChE J.* **63,** 532–550 (2017).
16. Rogers, A. & Ierapetritou, M. *12th International Symposium on Process Systems Engineering and 25th European Symposium on Computer Aided Process Engineering*. *Computer Aided Chemical Engineering* **37,** (Elsevier, 2015).
17. Singh, R., Sen, M., Ierapetritou, M. & Ramachandran, R. Integrated Moving Horizon-Based Dynamic Real-Time Optimisation and Hybrid MPC-PID Control of a Direct Compaction Continuous Tablet Manufacturing Process. *J. Pharm. Innov.* **10,** 233–253 (2015).
18. Abejón, R., Garea, A. & Irabien, A. Analysis and optimisation of continuous organic solvent nanofiltration by membrane cascade for pharmaceutical separation. *AIChE J.* **60,** 931–948 (2014).
19. Abejón, R., Garea, A. & Irabien, A. Organic Solvent Recovery and Reuse in Pharmaceutical Purification Processes by Nanofiltration Membrane Cascades. *Chem. Eng. Trans.* **43,** (2015).
20. Sahlodin, A. M. & Barton, P. I. Optimal Campaign Continuous Manufacturing. *Ind. Eng. Chem. Res.* **54,** 11344–11359 (2015).
21. Ott, D. *et al.* Life Cycle Analysis within Pharmaceutical Process Optimisation and Intensification: Case Study of Active Pharmaceutical Ingredient Production. *ChemSusChem* **7,** 3521–3533 (2014).
22. Ott, D., Borukhova, S. & Hessel, V. Life cycle assessment of multi-step rufinamide synthesis – from isolated reactions in batch to continuous microreactor networks. *Green Chem.* **18,** 1096–1116 (2016).
23. Kralisch, D. *et al.* Rules and benefits of Life Cycle Assessment in green chemical process and synthesis design: a tutorial review. *Green Chem.* **17,** 123–145 (2015).
24. Jolliffe, H. G. & Gerogiorgis, D. I. Plantwide Design and Evaluation of Two Continuous Pharmaceutical Manufacturing (CPM) Cases: Ibuprofen and Artemsinin. *Comput. Aided Chem. Eng.* (2016). doi:10.1016/B978-0-444-63576-1.50063-7
25. Jolliffe, H. G. & Gerogiorgis, D. I. Technoeconomic Optimisation of a Conceptual Flowsheet for Continuous Separation of an Analgaesic Active Pharmaceutical Ingredient (API). *Ind. Eng. Chem. Res.* **56,** 4357–4376 (2017).
26. Horváth, Z. *et al.* Recovery of Artemisinin from a Complex Reaction Mixture Using Continuous Chromatography and Crystallisation. (2015).
27. Malwade, C. R. *et al.* Crystallisation of Artemisinin from Chromatography Fractions of *Artemisia annua* Extract. *Org. Process Res. Dev.* **20,** 646–652 (2016).
28. Jolliffe, H. G. & Gerogiorgis, D. I. Technoeconomic optimisation and comparative environmental impact evaluation of continuous crystallisation and antisolvent selection for artemisinin recovery. *Comput. Chem. Eng.* **103,** 218–232 (2017).
29. Gracin, S., Brinck, T. & Rasmuson, A. C. Prediction of Solubility of Solid Organic Compounds in Solvents by UNIFAC. *Ind. Eng. Chem. Res.* **41,** 5114–5124 (2002).
30. Bouillot, B., Teychene, S. & Biscans, B. An Evaluation of Thermodynamic Models for the Prediction of Drug and Drug-like Molecule Solubility in Rrganic Solvents. *Fluid Phase Equilib.* **309,** 36–52 (2011).
31. Renon, H. & Prausnitz, J. M. Local Compositions in Thermodynamic Excess Functions for Liquid Mixtures. *AIChE J.* **14,** 135–144 (1968).
32. Abrams, D. S. & Prausnitz, J. M. Statistical thermodynamics of liquid-mixtures - new expression for excess Gibbs energy of partly or completely miscible systems. *AIChE J.* **21,** 116–128 (1975).

33. Chen, C. C. & Crafts, P. A. Correlation and prediction of drug molecule solubility in mixed solvent systems with the Nonrandom Two-Liquid Segment Activity Coefficient (NRTL-SAC) model. *Ind. Eng. Chem. Res.* **45,** 4816–4824 (2006).
34. Nti-Gyabaah, J., Gbewonyo, K. & Chiew, Y. C. Solubility of artemisinin in different single and binary solvent mixtures between (284.15 and 323.15) K and NRTL interaction parameters. *J. Chem. Eng. Data* **55,** 3356–3363 (2010).
35. Carta, R., Dernini, S. & Sanna, P. Isobaric vapor-liquid equilibria for the ternary system acetone-ethyl acetate-ethanol. *J. Chem. Eng. Data* **29,** 463–466 (1984).
36. Alder, C. M. *et al.* Updating and further expanding GSK's solvent sustainability guide. *Green Chem.* **18,** 3879–3890 (2016).
37. Schaber, S. D. *et al.* Economic Analysis of Integrated Continuous and Batch Pharmaceutical Manufacturing: A Case Study. *Ind. Eng. Chem. Res.* **50,** 10083–10092 (2011).
38. Woods, D. R. *Rules of Thumb in Engineering Practice*. (Wiley, 2007).
39. Powell, M. J. D. The BOBYQA algorithm for bound constrained optimisation without derivatives. (2009).
40. Audet, C. & Dennis, J. E. Mesh Adaptive Direct Search Algorithms for Constrained Optimisation. *SIAM J. Optim.* **17,** 188–217 (2006).
41. Le Digabel, S. & Sébastien. Algorithm 909. *ACM Trans. Math. Softw.* **37,** 1–15 (2011).
42. *Computational Pharmaceutical Solid State Chemistry*. (John Wiley & Sons, Inc, 2016). doi:10.1002/9781118700686
43. Sheldon, R. A. Fundamentals of Green Chemistry: Efficiency in Reaction Design. *Chem. Soc. Rev.* **41,** 1437–1451 (2012).
44. Ritter, S. K. Reducing Environmental Impact of Organic Synthesis. *Chem. Eng. News* **91,** 22–23 (2013).

APPENDIX A: SOLVER OPTIMISATION PATHS

Figure A1 illustrates the optimisation paths followed for various starting points (x_0) when implementing NLOPT-BOBYQA. Optimisation paths are shorter (fewer iterations are required) when initial points are closer to the solution; the paths are particularly long when initial m_{AS} and $T_{Crys,LO}$ guesses are higher (>65% and >20 °C, respectively) . As the solver approaches the optimal solution, the difference between guesses become smaller.

Figure A2 shows the optimisation paths followed by the NOMAD solver for various starting points (x_0). While the NOMAD solver frequently changes direction and explores a majority of the design space, it nevertheless also reliably finds the same optimal point for all initial guess. An initial guess closer to the optimal point reduces the solution time and decreases the number of iterations required, and decreases the range of the design space that the algorithm explores before finding the optimal point.

The NLOPT-BOBYQA solver explores a significantly reduced design space range (Figure A1) compared to NOMAD, doubling back on itself in far fewer instances. It also reaches a solution faster than NOMAD, while achieving the same level of precision. As NLOPT-BOBYQA solves much faster than NOMAD without significant differences in precision, it was chosen as the main solver. Except for the solution time, there was no difference from using different initial guesses, and so for most cases an initial guess of x_0 = (55, 50, 10), was used.

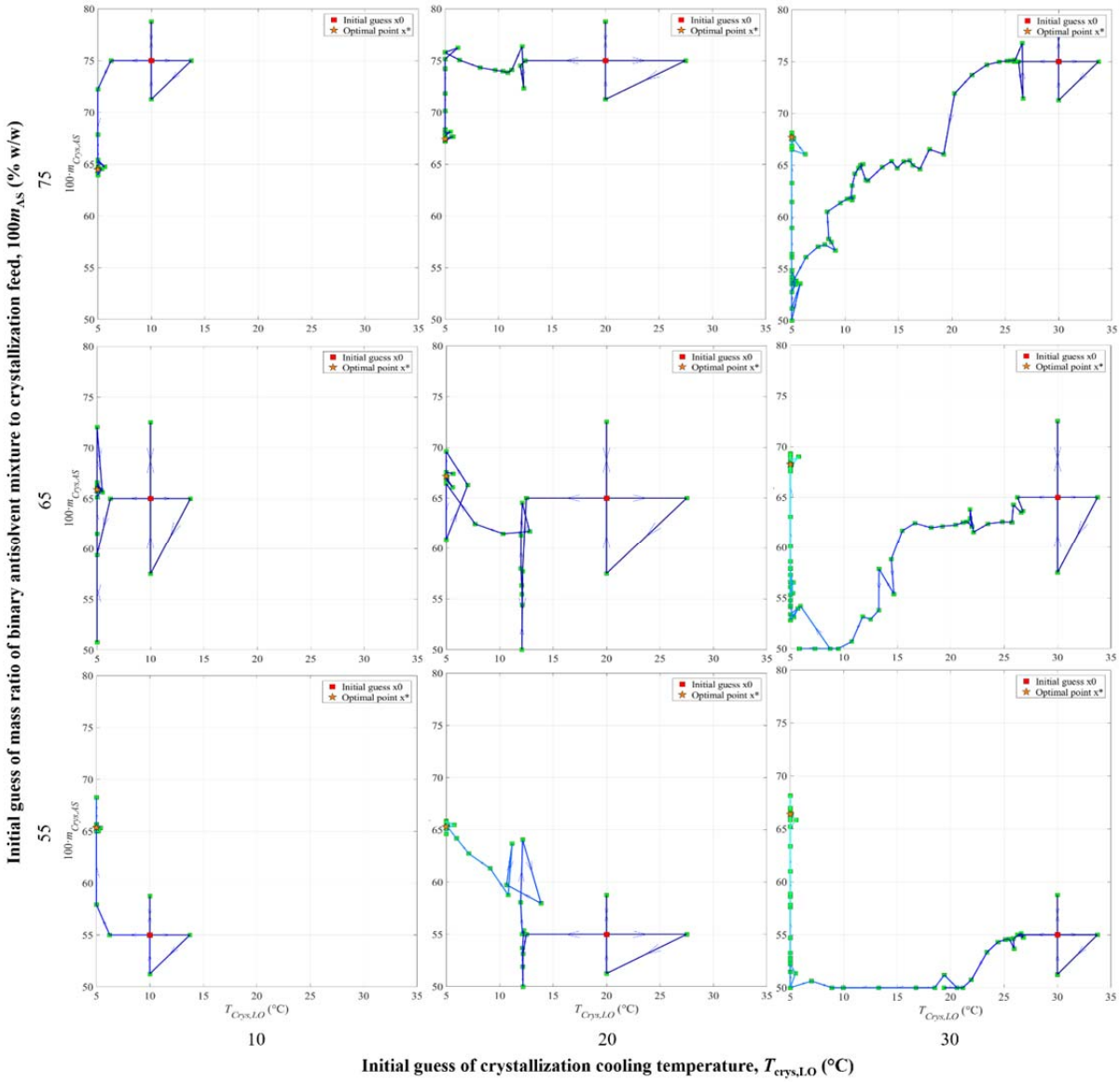

Figure A1. Typical NLOPT-BOBYQA optimisation path followed; multiple initial guesses (x_0) shown. Acetone-ethyl acetate antisolvent stage for one crystallizer.

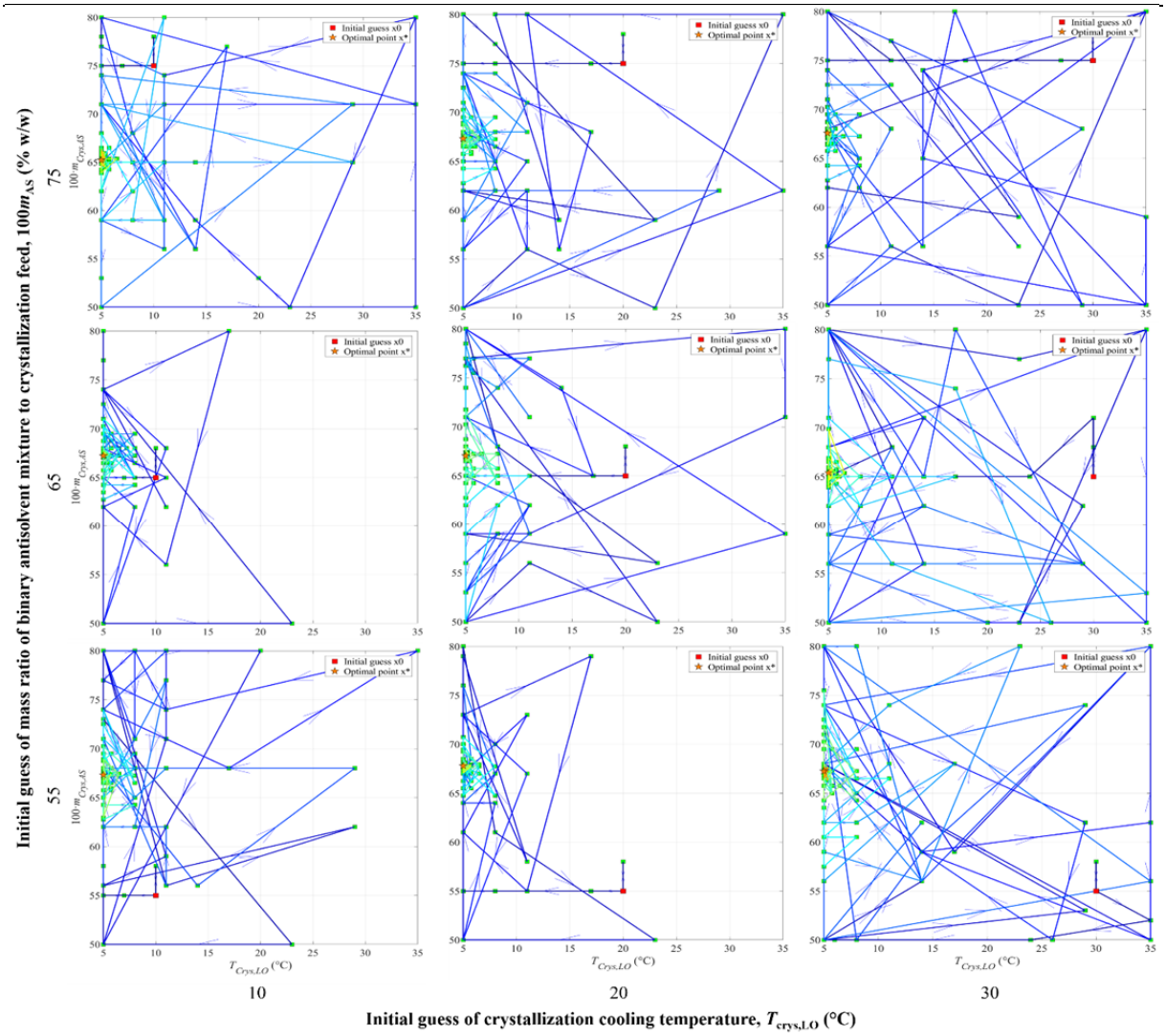

Figure A2. Typical NOMAD optimisation path followed; multiple initial guesses (x_0) shown. Acetone-ethyl acetate antisolvent stage for one crystallizer.

APPENDIX B: TOTAL COST RESPONSE SURFACES

Figures B1 and B2 show total cost response surfaces of the design spaces for fixed and variable size crystallizers, implementing 1-3 crystallizers, for all binary antisolvent mixtures considered. Values plotted in Figure 3 correspond to corners of the response surfaces in Figures B1 and B2 and total cost minima found in these design spaces via the nonlinear optimisation procedure described in this work.

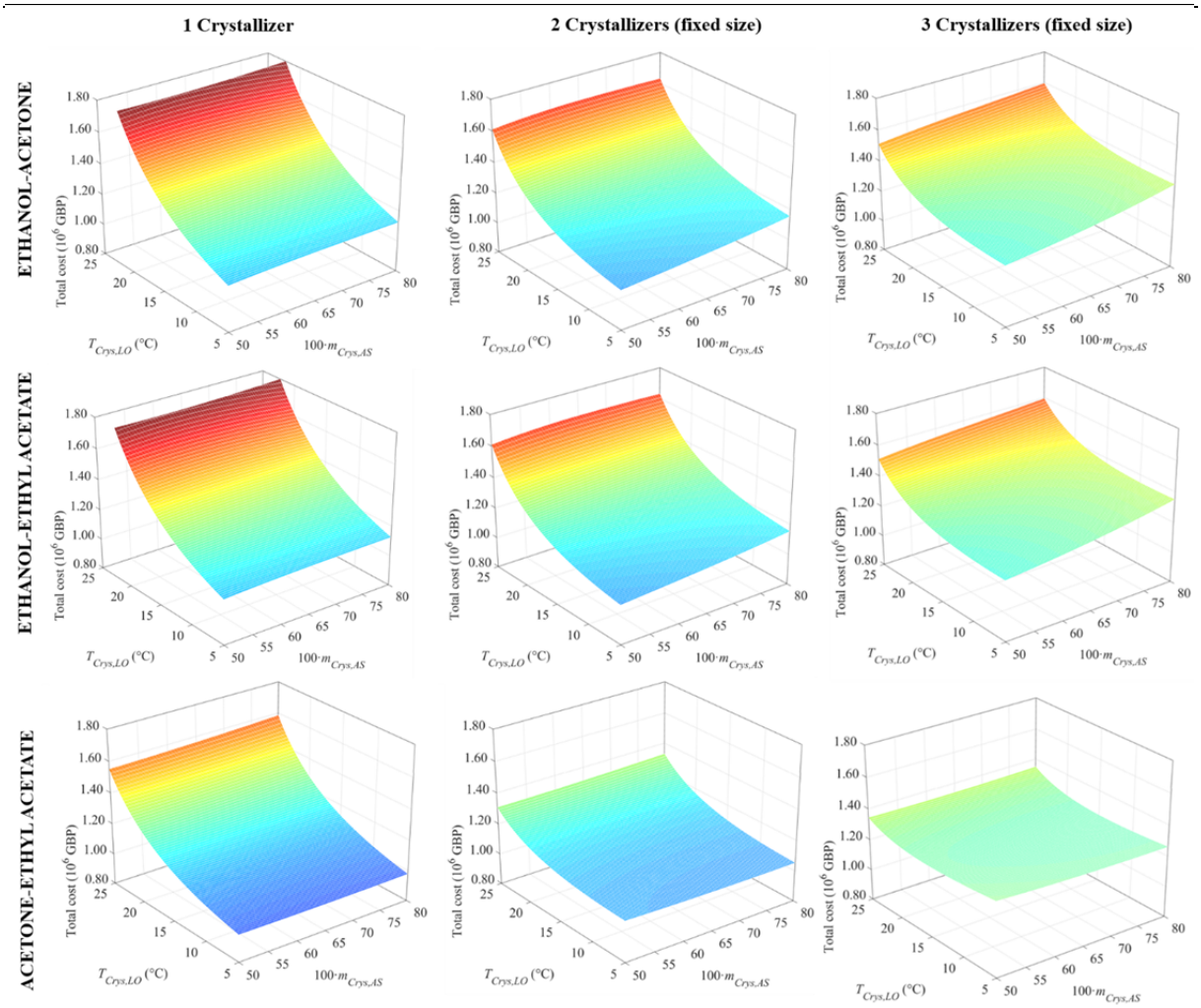

Figure B1. Total cost response surfaces for fixed-size crystallizers. Optimal ethanol–acetone antisolvent mix (100:0); optimal ethanol–ethyl acetate antisolvent mix (99.9: 0.1). C) Optimal acetone – ethyl acetate antisolvent mix (2:98).

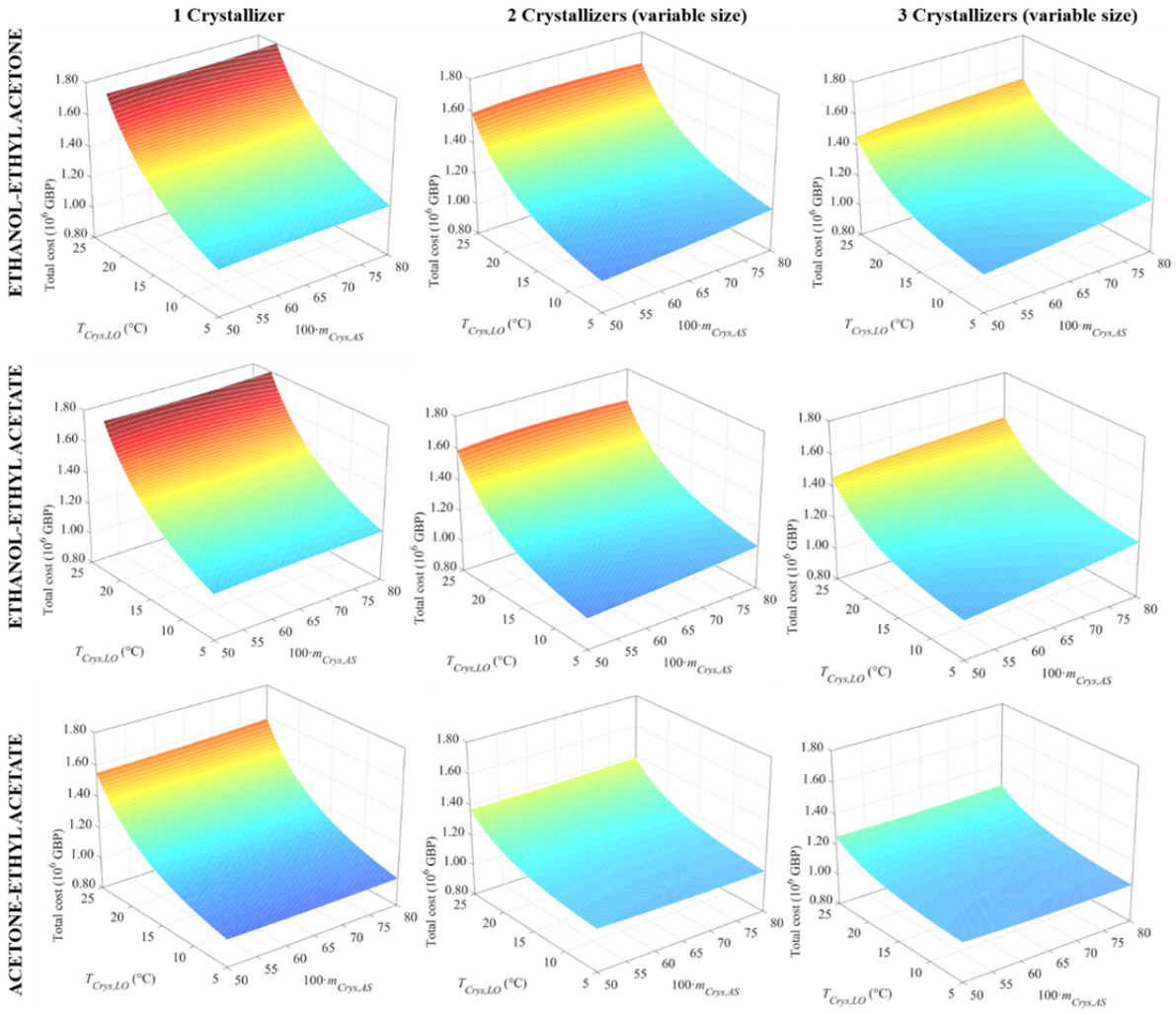

Figure B2. Total cost response surfaces for variable-size crystallizers. Optimal ethanol–acetone antisolvent mix (100:0); optimal ethanol–ethyl acetate antisolvent mix (99.9: 0.1). C) Optimal acetone – ethyl acetate antisolvent mix (2:98).